RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/005,318F

DATE: 08/29/2003
TIME: 09:41:28

Input Set : A:\Epi3004b.app
Output Set: N:\CRF4\08292003\I005318F.raw

```
 3 <110> APPLICANT: HEIN, MICH B.
 4        HIATT, ANDREW C.
 5        FITCHEN, JOHN H.
 7 <120> TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
 9 <130> FILE REFERENCE: EPI3004B
11 <140> CURRENT APPLICATION NUMBER: 09/005,318F
12 <141> CURRENT FILING DATE: 1998-01-09
14 <150> PRIOR APPLICATION NUMBER: 08/782,481
15 <151> PRIOR FILING DATE: 1997-01-10
17 <150> PRIOR APPLICATION NUMBER: 09/005,167
18 <151> PRIOR FILING DATE: 1998-01-09
20 <160> NUMBER OF SEQ ID NOS: 113
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 137
26 <212> TYPE: PRT
27 <213> ORGANISM: Homo sapiens
29 <400> SEQUENCE: 1
30 Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys Ala
31   1               5                  10                  15
33 Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu Asp
34                  20                  25                  30
36 Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg Glu
37              35                  40                  45
39 Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Pro Val Tyr His
40          50                  55                  60
42 Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu Asp
43   65                  70                  75                  80
45 Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp Ser
46                  85                  90                  95
48 Ala Thr Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Ala
49              100                 105                 110
51 Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr Ala
52          115                 120                 125
54 Leu Thr Pro Asp Ala Cys Tyr Pro Asp
55      130                 135
58 <210> SEQ ID NO: 2
59 <211> LENGTH: 135
60 <212> TYPE: PRT
61 <213> ORGANISM: Mus sp.
63 <400> SEQUENCE: 2
64 Gln Asp Glu Asn Glu Arg Ile Val Val Asp Asn Lys Cys Lys Cys Ala
65   1               5                  10                  15
```

RAW SEQUENCE LISTING   DATE: 08/29/2003
PATENT APPLICATION: US/09/005,318F   TIME: 09:41:28

Input Set : A:\Epi3004b.app
Output Set: N:\CRF4\08292003\I005318F.raw

```
67 Arg Ile Thr Ser Arg Ile Ile Pro Ser Ala Glu Asp Pro Ser Gln Asp
68                 20                  25        .         30
70 Ile Val Glu Arg Asn Val Arg Ile Ile Val Pro Leu Asn Ser Arg Glu
71             35                  40                  45
73 Asn Ile Ser Asp Pro Thr Ser Pro Met Arg Thr Lys Pro Val Tyr His
74       50                  55                  60
76 Leu Ser Asp Leu Cys Lys Lys Cys Asp Thr Thr Glu Val Glu Leu Glu
77 65                  70                  75                  80
79 Asp Gln Val Val Thr Ala Ser Gln Ser Asn Ile Cys Asp Ser Asp Ala
80                 85                  90                  95
82 Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Asn Arg Val
83             100                 105                 110
85 Lys Leu Ser Tyr Arg Gly Gln Thr Lys Met Val Glu Thr Ala Leu Thr
86             115                 120                 125
88 Pro Asp Ser Cys Tyr Pro Asp
89       130                 135
92 <210> SEQ ID NO: 3
93 <211> LENGTH: 137
94 <212> TYPE: PRT
95 <213> ORGANISM: Oryctolagus cuniculus
97 <400> SEQUENCE: 3
98 Asp Asp Glu Ala Thr Ile Leu Ala Asp Asn Lys Cys Met Cys Thr Arg
99 1                 5                  10                  15
101 Val Thr Ser Arg Ile Ile Pro Ser Thr Glu Asp Pro Asn Glu Asp Ile
102                 20                  25                  30
104 Val Glu Arg Asn Ile Arg Ile Val Val Pro Leu Asn Asn Arg Glu Asn
105             35                  40                  45
107 Ile Ser Asp Pro Thr Ser Pro Leu Arg Arg Asn Pro Val Tyr His Leu
108       50                  55                  60
110 Ser Asp Val Cys Lys Lys Cys Asp Pro Val Glu Val Glu Leu Glu Asp
111 65                  70                  75                  80
113 Gln Val Val Thr Ala Thr Gln Ser Asn Ile Cys Asn Glu Asp Asp Gly
114                 85                  90                  95
116 Val Pro Glu Thr Cys Tyr Met Tyr Asp Arg Asn Lys Cys Tyr Thr Thr
117             100                 105                 110
119 Met Val Pro Leu Arg Tyr His Gly Glu Thr Lys Met Val Gln Ala Ala
120             115                 120                 125
122 Leu Thr Pro Asp Ser Cys Tyr Pro Asp
123       130                 135
126 <210> SEQ ID NO: 4
127 <211> LENGTH: 136
128 <212> TYPE: PRT
129 <213> ORGANISM: Bos sp.
131 <400> SEQUENCE: 4
132 Glu Asp Glu Ser Thr Val Leu Val Asp Asn Lys Cys Gln Cys Val Arg
133 1                 5                  10                  15
135 Ile Thr Ser Arg Ile Ile Arg Asp Pro Asp Asn Pro Ser Glu Asp Ile
136             20                  25                  30
138 Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Thr Arg Glu Asn
```

RAW SEQUENCE LISTING   DATE: 08/29/2003
PATENT APPLICATION: US/09/005,318F   TIME: 09:41:28

Input Set : A:\Epi3004b.app
Output Set: N:\CRF4\08292003\I005318F.raw

```
139          35                    40                    45
141 Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Glu Pro Lys Tyr Asn Leu
142     50                    55                    60
144 Ala Asn Leu Cys Lys Lys Cys Asp Pro Thr Glu Ile Glu Leu Asp Asn
145  65                    70                    75                    80
147 Gln Val Phe Thr Ala Ser Gln Ser Asn Ile Cys Pro Asp Asp Asp Tyr
148                    85                    90                    95
150 Ser Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Thr Leu
151             100                   105                   110
153 Val Pro Ile Thr His Arg Gly Val Thr Arg Met Val Lys Ala Thr Leu
154          115                   120                   125
156 Thr Pro Asp Ser Cys Tyr Pro Asp
157     130                   135
160 <210> SEQ ID NO: 5
161 <211> LENGTH: 119
162 <212> TYPE: PRT
163 <213> ORGANISM: Rana sp.
165 <220> FEATURE:
166 <221> NAME/KEY: MOD_RES
167 <222> LOCATION: (47)
168 <223> OTHER INFORMATION: Variable amino acid
170 <220> FEATURE:
171 <221> NAME/KEY: MOD_RES
172 <222> LOCATION: (88)...(89)
173 <223> OTHER INFORMATION: Variable amino acid
175 <220> FEATURE:
176 <221> NAME/KEY: MOD_RES
177 <222> LOCATION: (91)
178 <223> OTHER INFORMATION: Variable amino acid
180 <400> SEQUENCE: 5
181 Glu Gln Glu Tyr Ile Leu Ala Asn Asn Lys Cys Lys Cys Val Lys Ile
182   1               5                    10                    15
184 Ser Ser Arg Phe Val Pro Ser Thr Glu Arg Pro Gly Glu Glu Ile Leu
185             20                    25                    30
```
W-->
```
187 Glu Arg Asn Ile Gln Ile Thr Ile Pro Thr Ser Ser Arg Met Xaa Ile
188          35                    40                    45
190 Ser Asp Pro Tyr Ser Pro Leu Arg Thr Gln Pro Val Tyr Asn Leu Trp
191     50                    55                    60
193 Asp Ile Cys Gln Lys Cys Asp Pro Val Gln Leu Glu Ile Gly Gly Ile
194  65                    70                    75                    80
196 Pro Val Leu Ala Ser Gln Pro Xaa Xaa Ser Xaa Pro Asp Asp Glu Cys
197                    85                    90                    95
199 Tyr Thr Thr Glu Val Asn Phe Lys Lys Lys Val Pro Leu Thr Pro Asp
200             100                   105                   110
202 Ser Cys Tyr Glu Tyr Ser Glu
203          115
206 <210> SEQ ID NO: 6
207 <211> LENGTH: 128
208 <212> TYPE: PRT
```

RAW SEQUENCE LISTING   DATE: 08/29/2003
PATENT APPLICATION: US/09/005,318F   TIME: 09:41:28

Input Set : A:\Epi3004b.app
Output Set: N:\CRF4\08292003\I005318F.raw

```
209 <213> ORGANISM: Lumbricus sp.
211 <400> SEQUENCE: 6
212 Asn Lys Cys Met Cys Thr Arg Val Thr Ala Arg Ile Arg Gly Thr Arg
213   1               5                  10                  15
215 Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Tyr Ile Arg Ile Asn Val
216               20                  25                  30
218 Pro Leu Lys Asn Arg Gly Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg
219           35                  40                  45
221 Asn Gln Pro Val Tyr His Leu Ser Pro Ser Cys Lys Lys Cys Asp Pro
222       50                  55                  60
224 Tyr Glu Asp Gly Val Val Thr Ala Thr Glu Thr Asn Ile Cys Tyr Pro
225  65                  70                  75                  80
227 Asp Gln Gly Val Pro Gln Ser Cys Arg Asp Tyr Cys Pro Glu Leu Asp
228                   85                  90                  95
230 Arg Asn Lys Cys Tyr Thr Val Leu Val Pro Pro Gly Tyr Thr Gly Glu
231               100                 105                 110
233 Thr Lys Met Val Gln Asn Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
234           115                 120                 125
237 <210> SEQ ID NO: 7
238 <211> LENGTH: 421
239 <212> TYPE: DNA
240 <213> ORGANISM: Homo sapiens
242 <220> FEATURE:
243 <221> NAME/KEY: CDS
244 <222> LOCATION: (1)..(414)
246 <220> FEATURE:
247 <221> NAME/KEY: sig_peptide
248 <222> LOCATION: (1)..(6)
250 <220> FEATURE:
251 <221> NAME/KEY: mat_peptide
252 <222> LOCATION: (7)..(414)
254 <400> SEQUENCE: 7
255 gat cag gaa gat gaa cgt att gtt ctg gtt gac aac aag tgc aag tgt    48
256 Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys
257      -1   1               5                  10
259 gct cgt att act tct aga atc atc cgt agc tca gag gac cca aat gaa    96
260 Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu
261  15                  20                  25                  30
263 gat ata gtc gaa cgt aac atc cgt atc atc gtc cca ctg aat aac cgg   144
264 Asp Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg
265               35                  40                  45
267 gag aat atc tca gat cct aca agt ccg ttg cgc aca cgc ttc gta tac   192
268 Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Phe Val Tyr
269               50                  55                  60
271 cac ctg tca gat ctg tgt aag aag tgt gat cca aca gag gta gag ctg   240
272 His Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu
273               65                  70                  75
275 gac aat cag ata gtc act gcg act caa agc aac att tgc gat gag gac   288
276 Asp Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp
```

```
277      80                    85                    90
279 agc gct aca gaa acc tgc agc acc tac gat agg aac aaa tgc tac acg    336
280 Ser Ala Thr Glu Thr Cys Ser Thr Tyr Asp Arg Asn Lys Cys Tyr Thr
281  95              100                   105                   110
283 gcc gtg gtt ccg ctc gtg tat ggt gga gag aca aaa atg gtg gaa act    384
284 Ala Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr
285                 115                   120                   125
287 gcc ctt acg ccc gat gca tgc tat ccg gac tgaattc                    421
288 Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
289             130                   135
293 <210> SEQ ID NO: 8
294 <211> LENGTH: 215
295 <212> TYPE: DNA
296 <213> ORGANISM: Homo sapiens
298 <220> FEATURE:
299 <221> NAME/KEY: CDS
300 <222> LOCATION: (1)..(213)
302 <400> SEQUENCE: 8
303 gat cag aag tgc aag tgt gct cgt att act tct aga atc atc cgt agc     48
304 Asp Gln Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser
305   1           5                   10                  15
307 tca gag gac cca aat gaa gat ata gtc gaa cgt aac atc cgt atc atc     96
308 Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
309             20                  25                  30
311 gtc cca ctg aat aac cgg gag aat atc tca gat cct aca agt ccg ttg    144
312 Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
313         35                  40                  45
315 cgc aca cgc ttc gta tac cac ctg tca gat ctg tgt aag aag gat gag    192
316 Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Asp Glu
317      50                  55                  60
319 gac agc gct aca gaa acc tgc tg                                     215
320 Asp Ser Ala Thr Glu Thr Cys
321  65                  70
325 <210> SEQ ID NO: 9
326 <211> LENGTH: 140
327 <212> TYPE: DNA
328 <213> ORGANISM: Homo sapiens
330 <400> SEQUENCE: 9
331 ctagaatcat ccgtagctca gaggacccaa atgaagatat agtcgaacgt aacatccgta 60
332 tcatcgtccc actgaataac cgggagaata tctcagatcc tacaagtccg ttgcgcacac 120
333 gcttcgtata ccacctgtca                                              140
336 <210> SEQ ID NO: 10
337 <211> LENGTH: 31
338 <212> TYPE: DNA
339 <213> ORGANISM: Homo sapiens
341 <400> SEQUENCE: 10
342 gatcagaagt gcaagtgtgc tcgtattact t                                  31
345 <210> SEQ ID NO: 11
346 <211> LENGTH: 44
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 08/29/2003
PATENT APPLICATION: US/09/005,318F   TIME: 09:41:29

    Input Set : A:\Epi3004b.app
    Output Set: N:\CRF4\08292003\I005318F.raw

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; Xaa Pos. 47,88,89,91